



Zerto Scale and Benchmarking Guidelines

Rev01

Dec 2020

ZVR-ZSB-8.0

© 2020 Zerto All rights reserved.

Information in this document is confidential and subject to change without notice and does not represent a commitment on the part of Zerto Ltd. Zerto Ltd. does not assume responsibility for any printing errors that may appear in this document. No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or information storage and retrieval systems, for any purpose other than the purchaser's personal use, without the prior written permission of Zerto Ltd. All other marks and names mentioned herein may be trademarks of their respective companies.

The scripts are provided by example only and are not supported under any Zerto support program or service. All examples and scripts are provided "as-is" without warranty of any kind. The author and Zerto further disclaim all implied warranties including, without limitation, any implied warranties of merchantability or of fitness for a particular purpose.

In no event shall Zerto, its authors, or anyone else involved in the creation, production, or delivery of the scripts be liable for any damages whatsoever (including, without limitation, damages for loss of business profits, business interruption, loss of business information, or other pecuniary loss) arising out of the use of or inability to use the sample scripts or documentation, even if the author or Zerto has been advised of the possibility of such damages. The entire risk arising out of the use or performance of the sample scripts and documentation remains with you.

ZVR-ZSB-8.0

Zerto Scale and Benchmarking Guidelines

This document provides the user with an understanding of the performance and scale capabilities of Zerto Virtual Replication. It presents best practices regarding performance and scale considerations when designing a Zerto Virtual Replication solution. The document should not be used to determine precise sizing of a specific configuration. For assistance with final system specifications to meet your requirements, consult your Zerto Account team. Users should be familiar with Zerto Virtual Replication and have basic knowledge of storage technologies and virtualization technologies.

The following guidelines are provided:

- [Zerto Performance Fundamentals on page 4](#)
- [Zerto Sizing Limits on page 15](#)
- [Zerto Benchmarking Test Results on page 17](#)
- [Estimating WAN Bandwidth for VMware and Hyper-V on page 24](#)

Related Documentation:

- Installation Guides
- Administration Guides
- [Using the Zerto WAN Estimator](#)
- Enterprise Guidelines (Prerequisites and Requirements)
- [Release Notes](#)

Zerto Performance Fundamentals

- [Protected VMs on page 4](#)
- [Application IOs on page 4](#)
- [What is a VPG? on page 4](#)
- [Disk Considerations on page 5](#)
- [VRA Considerations on page 5](#)
- [ZVM Considerations on page 5](#)
- [ZVM Database Considerations on page 6](#)
- [Recommendations and Considerations When Using an External SQL Server on page 7](#)
- [ZCC Considerations on page 7](#)
- [Long-term Retention Considerations on page 8](#)
- [Public Cloud Considerations on page 10](#)
- [Networking Considerations on page 11](#)
- [WAN Optimization Considerations on page 11](#)
- [Journal Considerations on page 14](#)
- [Best Practices for Multiple VRA Pairings on page 14](#)

Protected VMs

Protected VMs are the VMs on the production site that are being protected to the recovery site by Zerto Virtual Replication. The Zerto Virtual Manager on the production site is paired with a ZVM (or ZVM in a ZCA) on the recovery site.

Application IOs

The VRA only replicates application writes from its host's protected VMs to the recovery hosts and its peer VRA or ZCA. To estimate the performance of a VRA's replication load, only the application writes to paired VRAs need to be considered. On the recovery site VRA, the performance is estimated from a variety of loads including but not limited to all the replication streams from peer VRAs plus the journal writes.

What is a VPG?

A VPG is a logical grouping of VMs and their associated disks that are replicated together in a way that ensures write-order consistency. The VPG also contains settings, such as boot order and re-IP details, that are used in the recovery operations.

Disk Considerations

Each VRA can manage a set maximum of volumes, whether these are volumes being protected or recovered.

You can only protect a virtual machine in a VPG when the virtual machine has no more than 60 disks.

- In VMware vSphere environments, this requires 4 SCSI controllers each with a maximum of 15 disks.
- In Microsoft Hyper-V environments, this can be a combination of IDE and SCSI disks, where the total of IDE and SCSI disks must not exceed 60 disks. Each virtual machine can have up to:
 - 2 IDE controllers each with a maximum of 4 IDE disks
 - 4 SCSI controllers each with a maximum of 15 disks

VRA Considerations

Note: This section only applies to non-Zerto versions.

A VRA is a Zerto Virtual Replication appliance that manages the replication of virtual machines locally or across sites. A VRA must be installed on every hypervisor host with virtual machines that require protecting in the protected site and on every host to which VMs may recover to in the recovery site. The VRA compresses the data that is passed across the WAN from the protected site to the recovery site.

Each VRA can be configured with 1 or 2 vCPUs.

The VRA memory allocation can range from 1GB to 16GB. The amount of memory allocated to a VRA when it is installed determines the maximum buffer size for the VRA for buffering IOs written by the protected virtual machines before the writes are sent over the network to the recovery VRA. The recovery VRA also buffers the incoming IOs until they are written to the journal. If a buffer becomes full, a Bitmap Sync is performed after space is freed up in the buffer.

Depending on the replication rate from the production site VRA, you can estimate the number of vCPUs and memory. If the production site VRA is replicating to multiple VRAs or replicating multiple VMs with a high change rate (over 100MBPS), then consider allocating more vCPUs and memory to the VRA.

On the recovery site VRA, the number of vCPUs and memory is like the production site VRA except the allocation needs to accommodate the replication data being written to the mirror disks, journal updates and checkpoints. If there are multiple production VRAs replicating to a single recovery site VRA and generating high incoming traffic to the recovery VRA, then consider allocating more vCPUs and memory to the recovery VRA.

ZVM Considerations

Each ZVM monitors the virtualization management system such as VMware vCenter and Microsoft System Center for status of the protected VMs.

ZVM performance can be impacted by:

- Number of VPGs managed by the ZVM
- Ongoing error states & alerts
- ZVM or VRA restarts & intermediate disconnection issues
- Number of peers & VRA
- Use of APIs

ZVM Database Considerations

Zerto Virtual Replication requires a database to manage information for disaster recovery.

Zerto requires a database to manage data for scalable business continuity software solutions.

- During the Zerto Virtual Manager installation, the user is able to select whether to install and use an **embedded** SQL Server (**localdb**) as the database.
- Alternatively, and also during the installation, the user is able to choose whether to instead select and use an **external** SQL Server instance.
- The larger the environment protected by Zerto Virtual Manager, the larger the database size required to support it.
- When using a **localdb**, the size of the environment that can be protected by Zerto Virtual Manager, is limited.
- **Failure to follow the sizing guideline can result in software errors, and not just performance degradation.**
- After installing Zerto Virtual Manager, you are able to migrate from using an embedded SQL Server, to using an external SQL Server. For guidelines, considerations and procedures, see [Migrating the Zerto Virtual Replication Database to Microsoft SQL Server](#).
- For a list of supported external SQL server databases, see [Interoperability Matrix](#).
- Zerto recommends placing external SQL server on a separate VM from where the Zerto Virtual Manager is running.

Note: If SQL Server is used, it is your responsibility to make sure that database downtime is planned in coordination with your disaster recovery and business continuity requirements. During database downtime, there will be inconsistencies between the Zerto Virtual Managers, such as the management of checkpoints.

Use the following table to determine the deployment size and thus, the database to use for Zerto.

The criteria used to determine the deployment size are:

- Sites
- VPGs

When the deployment sizes are different across the criteria, always choose the database based on the **larger** deployment size.

Database Type	Max number of VPGs Which Can Replicate to the Current/Local Site	Max number of peer sites
Embedded SQL Server (localDB) / SQL Express Server	2,000	10
External SQL Server	10,000	Up to 100

Recommendations and Considerations When Using an External SQL Server

Before migrating your database, you must review this section.

Note: To perform a **database migration**, Zerto version must be **4.5Ux and later**. For details, refer to [Migrating the Zerto Virtual Replication Database to Microsoft SQL Server](#).

- Zerto recommends running the external database on a separate VM.
- Zerto recommends using faster storage for the database. For a Zerto site protecting more than 250 incoming VPG, we highly recommend using fast storage for the database (SSD).
- The SQL Server must use at least 4 CPUs. See the following table:

Number of VPG protecting to a site	SQL Server memory	Disk space
Up to 1000 incoming VPG	8 GB	10 GB
1000 - 2000 incoming VPG	16 GB	25 GB
2000 - 5000 incoming VPG	16 GB	50 GB
5000 - 10000 incoming VPG	16 GB	100 GB

ZCC Considerations

A Zerto Cloud Connector is used in multi-tenancy solutions in which disaster recovery services are offered to internal or external customers. The ZCC requires the following, regardless of the size of the site:

- 4GB disk space
- At least 1GB of reserved memory
- 2 vCPU
- 2 vNICs

Zerto recommends using a 10Gbps vNIC for each Zerto Cloud Connector, enabling it to handle 10Gbps of traffic.

Long-term Retention Considerations

The Long-term Retention feature provides longer term recovery from a repository by protecting the relevant virtual machines in Virtual Protection Groups.

When using Long-term Retention, additional considerations are needed for an optimized experience.

Repositories

Repository capacity estimation depends on several factors:

VM Size:	Total size of VMs being protected.
Frequency:	Frequency of retention processes such as dailies and weeklies.
Retention Period:	Time until the "restore points" will expire.
Data daily change rate:	Typical daily change rate is ~5%. This is subject to change per user, workload, and application etc.
Retention type:	Full or Incremental.
Data reduction ratios:	Deduplication and compression rates. This number varies between types of data, as well as between deduplicating storage/PBBA vendors. If the actual rates are not known, conservative estimates should be applied.
Capacity Utilization:	Considers spare capacity for cases like overcoming gaps in change rate estimations, change in Retention process' frequency, Manual Retention processes, or overlapping Retention sets (see the below examples).

Assumptions:

- (1*) In case there are no data reduction features activated on the storage level or the data itself is not considered as duplicated/compressible, used capacity is treated as equal to the allocated capacity.
- (2*) Metadata consumed capacity is negligible.
- (3*) For the first Daily Retention, a full Retention process will be performed which was not considered in the calculations below.

- (4*) A daily copy will be taken along with either weekly or monthly based on what the user has configured; this will be a full copy and can be deducted from the number of incremental daily copies.

Example 1:

Use case:

- 100 VMs, each with 1x250GB disks, 70% capacity utilized
- Daily change rate: 5%
- Frequency: 7 daily (incremental), 4 Weekly (full) ,12 Monthly (full) and 7 Yearly (full)
- No data reduction (1*)

Calculations:

- Full copy of the data: $(100 \times 250\text{GB used capacity}) \times (4 + 12 + 7 \text{ copies}) / 1024 = 561.5\text{TB}$
- Incremental copy of the data: $(100 \times 250\text{GB} \times 5\%(\text{change rate}) \times (7-1) \text{ dailies}) / 1024 = 8.5\text{TB}$
- Spare: $(100 \times 250\text{GB used capacity}) \times 2 = 49 \text{ TB}$
- Total : $561.5 + 8.5 + 49 + \sim 1\%(\text{metadata}) = \sim 625.2 \text{ TB}$

Example 2:

Use case:

- 100 VMs, each with 1x250GB disks, 70% capacity utilized
- Daily change rate: 5%
- Frequency: 7 daily (incremental), 4 Weekly (full) ,12 Monthly (full) and 7 Yearly (full)
- Data reduction rate: 1:4

Calculations:

- Full copy of the data: $((100 \times 175\text{GB used capacity}) \times 25\% \text{ Data Reduction}) \times (4 + 12 + 7 \text{ copies}) / 1024 = 98\text{TB}$
- Incremental copy of the data: $(100 \times 175\text{GB} \times 25\% \times 5\%(\text{change rate}) \times 7 \text{ dailies}) / 1024 = 1.5\text{TB}$
- Spare: $(100 \times 175\text{GB}) \times (25\% \text{ Data Reduction}) \times 2 = \sim 8.5\text{TB}$
- Total = $98 + 1.4 + 8.5 + \sim 1\%(\text{metadata}) = 109 \text{ TB}$

Indexing

There are several limitations to consider when using indexing:

- Storing the index for 100 restore points (incremental or full) will consume between 5-10% of the original VM size.
- Zerto guardrails the replication so that other processes will not be affected, including indexing.

- Indexing is done on the ZVM machine using a service called VBA.
- The VBA is limited to running 3 indexing tasks in parallel, and not more than one VRA.
- The impact on each VRA is similar to having a single VM mounted for File Level Restore.

Public Cloud Considerations

When protecting VMs in Azure and AWS, there are some additional considerations. The Zerto Virtual Replication solution with Azure and AWS includes deployment of one or more Zerto Cloud Appliances (ZCAs). The ZCA is a combination of a VRA and a ZVM that runs in an Azure or AWS instance in the public cloud.

When planning for a Public Cloud deployment, consider each major condition in the solution:

1. [Steady State Replication to the Public Cloud on page 10](#)
2. [Resource Availability in the Public Cloud on page 11](#)
3. [Failover Handling to Public Cloud on page 11](#)

Additionally, if the RPO is rising, you can consider scaling up the ZCA and scaling out the ZCA. When scaling up the ZCA, you upgrade to a higher performance instance from the Public Cloud in terms of CPUs, memory and bandwidth. When scaling out the ZCA, you add more instances of the ZCA and configure the VPGs to use the additional ZCAs.

Steady State Replication to the Public Cloud

For steady state replication to the Public Cloud the following factors impact performance:

- Replication traffic from the protected VMs. Please see the Public Cloud instance limitations on bandwidth:
 - Azure information: <https://docs.microsoft.com/en-us/azure/virtual-network/virtual-machine-network-throughput>
 - AWS bandwidth information: <https://aws.amazon.com/ec2/instance-types/#instance-type-matrix>
- Size of the protected disks. Please see the below section on Sizing Limits for additional information.
- WAN networking: direct connection from the protected site to the ZCA in the Public Cloud and adequate networking bandwidth for initial sync and replication.

We recommend monitoring the ZCA in the performance monitor in the Public Cloud instance to baseline and manage performance and scale. Key items to monitor are the CPU consumption and incoming/outgoing network throughput.

Resource Availability in the Public Cloud

The ZCA instance in Azure and AWS should have at least 4 cores and 8GiB of memory. Each ZCA Public Cloud account needs to have adequate CPU resources to handle a failover of all VPGs that are replicating and can recover to that ZCA. Usually the default number of CPUs allocated to each account is not enough to handle all the CPUs in the VMs/VPGs that have the Public Cloud for the recovery site. Please see the following details for the default vCPUs for Public Cloud:

- Azure: <https://docs.microsoft.com/en-us/azure/azure-subscription-service-limits>
- AWS: <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/instance-optimize-cpu.html>

Failover Handling to Public Cloud

When failing over to Public Cloud, the journal data is written to the recovery disk before the recovered VM is powered on. This process is called promotion. The promotion process is a significant part of the final RTO for the application.

There are a few best practices for optimizing the failover to the Public Cloud:

- Wait for the initial sync data to be applied from the journal to the mirror before doing any failover testing. Use the Journal History UI to check that the earliest recovery point is after the VPG was created.
- Reduce the configured history to the minimal value for the fastest RTO.

Networking Considerations

The connectivity between the production and recovery sites must have the bandwidth capacity to handle the replication data. The minimum dedicated bandwidth must be at least 5 Mb/sec. The latency between peered VRAs must be 120ms or less.

See the document [Using the Zerto WAN Sizing Estimator](#), for best practices in allocating enough bandwidth between ZVR sites.

Zerto Virtual Replication does not support replicating through NAT translated devices via port forwarding. For replication to occur through NAT translated devices, all replication traffic must be within a VPN.

WAN Optimization Considerations

There are some factors that can affect system optimization that should be considered in addition to estimating the networking bandwidth. Zerto Virtual Replication uses the maximum bandwidth available as well as optimization and internal compression to maintain low RPOs. This removes the need to use third-party WAN acceleration devices. However, Zerto Virtual Replication does work with third-party WAN acceleration devices.

Note: Zerto Virtual Replication WAN compression can be disabled for any VPG as required.

In addition, bandwidth can be throttled as necessary on each ZVM and QoS priorities can be specified for each VPG. If the WAN link becomes saturated, disconnected, or the data change rate is too high to maintain continuous replication, Zerto Virtual Replication enters a Bitmap sync mode. In this mode, VRA driver stores an index of which blocks have changed. When the condition causing the Bitmap sync is resolved, the latest changes are referenced from the bitmap, read from the source disk, and replicated to the recovery site, so that continuous replication may resume.

See the following sections:

- [Bandwidth Throttling to Temporarily Manage Bandwidth on page 12](#)
- [Using Swap Disks to Lessen the Bandwidth Load on page 13](#)
- [Preseeding to Initially Lessen the Bandwidth Load on page 13](#)
- [Setting the Priority for Using the WAN by Multiple VPGs on page 13](#)
- [Pausing Replication on page 13](#)

Bandwidth Throttling to Temporarily Manage Bandwidth

Zerto can be configured to throttle the bandwidth it uses during a specific time period. During this period, replication does not use more than the bandwidth selected. At the end of the period, bandwidth throttling ends.

> To configure time-based throttling:

1. In the Zerto user interface, select **Site Settings > Performance and Throttling**.

The Performance and Throttling window is displayed.

Bandwidth

Bandwidth Throttling [?]

☐ Limited

☒ Time-based Throttling [?]

☒ Limited

1 MB/sec

From 00 00

To 00 00

Show advanced settings...

2. Select **Time-based Throttling**.
3. To limit the MB/sec, select **Limited**, and enter a value.
4. To throttle the bandwidth during specific times, select a time period:
 - **From:** The hour and the minute to start the throttling, using a 24-hour clock.
 - **To:** The hour and the minute to end the throttling, using a 24-hour clock.

For Example: If you know that the bandwidth needs specific throttling during a certain period, during the daily peak transaction period you can override the throttling of the bandwidth for these specific times.

5. Click **APPLY** or **SAVE**.

Important:

IO Throttling values should be changed **only** in coordination with Zerto support.

Using Swap Disks to Lessen the Bandwidth Load

A swap disk is a virtual disk that Zerto replicates and then isolates from further changes. If temporary files on a protected virtual machine are moved to a swap disk, although the disk itself is replicated to the recovery site, changes to it are not replicated. Zerto performs an initial replication of the swap disk so that a file system is available to the guest operating system of the protected machine. Once the initial synchronization of this disk is completed, the disk will no longer participate in replication. This feature helps reduce replication traffic, journal size, and overall storage usage.

Swap disks are configured for the individual disk in the virtual machine in the VPG definition.

Preseeding to Initially Lessen the Bandwidth Load

Preseeding allows a user to enable Zerto to compare an existing virtual disk available at the recovery site to the protected volume instead of replicating the entire protected volume. The preseed disk must be the same size as the protected volume. When configured, these volumes will perform a delta sync. The delta sync will read the same blocks of the disks, generate and MD5 checksum from those blocks, and compare the checksums. Only the blocks containing different checksums will be replicated to the recovery disk. This process can reduce the time and bandwidth needed to synchronize the protected and recovery disks.

Setting the Priority for Using the WAN by Multiple VPGs

Zerto enables prioritizing replication for each VPG. The priorities are:

- High
- Medium
- Low

Prioritization is used when there is bandwidth contention between sites or when the load on the VRA is more than can be accommodated. It enables maintaining a low RPO on the most critical VPGs. As bandwidth becomes available, lower priority VPGs are synchronized and returned to a continuous protection state.

Pausing Replication

Zerto enables VPG replication to be paused, for example, during maintenance.

Journal Considerations

Journal considerations are different between on-premise recovery sites and Public Cloud recovery sites. Zerto's Public Cloud architecture requires applying journal data to the recovery disk to happen before the recovery VM can be powered-on. This behavior is different than the Hypervisor-based recovery sites where Zerto can instantly create a view of the recovery disk without applying any data. Consequently, public cloud RTO is longer than on-premise solutions.

For more details about Journal Sizing, please see the ZVR Journal Sizing and Best Practices document for initial sizing. Additionally, the Journal reports in Zerto Analytics can be used to monitor actual Journal usage by a VPG.

VMware and Hyper-V On-Premise Recovery Sites - Journal Sizing

For on-premise solutions, the general formula to calculate Journal size is:

Journal size = History_in_seconds*Average_throughput_of_all_VM_volumes*Average_compression_ratio

For example, if the history is 3 days (259,200 seconds) and:

- the average throughput is 10 MBPS and
- the data is 50% compressible on average

Then the journal size for this VM would be:

$259,200 * 10 * 0.5 = 1,296,000 \text{ MBPS}$ (~1266 GB)

Public Cloud Recovery Sites - Journal Sizing

Due to the difference in the architecture of Public Cloud solutions, configured history should be as short as possible to achieve the fastest RTOs.

Best Practices for Multiple VRA Pairings

When multiple protected source VRAs are peered to a single VRA on the recovery site, then you can consider using 2 CPUs on the recovery VRA.

Zerto Sizing Limits

The following table specifies the scale limitations for Zerto v8.0. Most of these limits are not enforced by the system, which means that a system with higher scale can be deployed; however, exceeding these limits may result in unpredictable system behavior, and is therefore not recommended.

Parameter	Limit	Limit
	VMware & Hyper-V Platform	Public Cloud
VMs per ZVM	VMware: 10,000 Hyper-V: 4,000	100
VPGs per ZVM	10,000	100
ZVMs per vCenter / System Center	1	NA
Max volumes per ZVM	VMware: 32,767 volumes Hyper-V: 4095 volumes	NA
VRAs per host	1	NA
Volumes per VRA	1500	250
Disks per VM	60	60
VMs per VRA	1500	120
Size of total protected disks per VRA	96TB	NA
Max size of a single disk being protected	96TB* *Contact Zerto Support for extending this limit.	Azure: 8TB* AWS: 16TB For maximum supported disk size considerations, see Release Notes > Known Issues > Azure .
Concurrent ZVM operations	50 VMs and 150 volumes	50 VMs and 20 volumes

Parameter	Limit	Limit
	VMware & Hyper-V Platform	Public Cloud
Max disk size for JFLR	32TB	Azure: 8TB* AWS: 1TB For maximum supported disk size considerations, see Release Notes > Known Issues > Azure .
Number of entries (files and folders) that can be indexed in a day	100 million	NA
LTR Indexing disk max	10TB	NA

Zerto Benchmarking Test Results

The following sections cover the benchmark testing configuration and results.

- [vSphere testing environment for ZVM benchmarks on page 17](#)
- [Hyper-V testing environment for ZVM benchmarks on page 17](#)
- [ZVM initialization and Pairing times on page 18](#)
- [ZVM GUI Response Times on page 21](#)
- [API Response Times on page 22](#)
- [Initial Synchronization Benchmarks on page 23](#)

vSphere testing environment for ZVM benchmarks

The vSphere benchmark testing environment consists of:

- Running Zerto v8.0, VMware vSphere 6.5, 6.0 and 5.5, Windows Server 2016
- 6000 VMs and 6000 VPGs configured
- Each host has:
 - 4 CPU with 16 GB memory
 - 200 GB storage space assigned
- VRAs configured with 1 CPU and 3GB memory
- The production and recovery ZVMs configured with:
 - 32 GB vRAM
 - 16 vCPUs
 - External ZVM database on SQL with:
 - SQL Server 2014 on Windows 2012 R2
 - 16GB vRAM
 - 12 vCPUs

Hyper-V testing environment for ZVM benchmarks

The Hyper-V benchmark testing environment consists of:

- Running Zerto v8.0, Windows Server 2016, Windows Server 2012
- 1300 VMs and 1300 VPGs configured

- Each host has:
 - 4 CPU with 16 GB memory
 - 200 GB storage space assigned
- VRAs configured with 1 CPU and 3GB memory
- ZVMs configured with:
 - 16 GB and 8 CPUs for production ZVM
 - 16GB and 12 CPUs for recovery ZVM
 - External ZVM database on SQL
 - recovery 16GB and 4 CPUs
 - production 8GB and 4 CPUs

ZVM initialization and Pairing times

These benchmarks measure the time it takes until the system resumes replication after restart.

[vSphere environments - ZVM initialization and Pairing times on page 18](#)

[Hyper-V environments - ZVM initialization and Pairing times on page 19](#)

[RTO Benchmarks on page 19](#)

vSphere environments - ZVM initialization and Pairing times

The following benchmarks cover the time in minutes to recover according to configurations and numbers in [vSphere testing environment for ZVM benchmarks on page 17](#):

1. Initialize production and recovery ZVMs
2. Connect the production and recovery ZVMs for the first time
3. Pair the production and recovery ZVMs with a VRA

vSphere Production Site - ZVM Benchmarks

	Zerto 8.0 benchmark
ZVM Restart	6:44
ZVM - ZVM synchronization time	5:10
ZVM - VRA synchronization time	2:05

vSphere Recovery Site - ZVM Benchmarks

	Zerto 8.0 benchmark
ZVM Restart	29:57
ZVM - ZVM synchronization time	28 :07
ZVM - VRA synchronization time	1:47

Hyper-V environments - ZVM initialization and Pairing times

The following benchmarks cover the time in minutes to recover according to configurations and numbers in [Hyper-V testing environment for ZVM benchmarks on page 17](#):

1. Initialize production and recovery ZVMs
2. Connect the production and recovery ZVMs for the first time
3. Pair the production and recovery ZVMs with a VRA

Hyper-V Production Site - ZVM Benchmarks

	Zerto 8.0 benchmark
ZVM Restart	1:30
ZVM - ZVM synchronization time	0:41
ZVM - VRA synchronization time	0:17

Hyper-V Recovery Site - ZVM Benchmarks

	Zerto 8.0 benchmark
ZVM Restart	1:21
ZVM - ZVM synchronization time	0:33
ZVM - VRA synchronization time	0:22

RTO Benchmarks

[RTO Benchmarks - vSphere Production and vSphere Recovery on page 20](#)

[RTO Benchmarks - Hyper-V Production and Hyper-V Recovery on page 20](#)

[RTO Benchmarks - VMware Cloud Director Production and VMware Cloud Director Recovery on page 21](#)

RTO Benchmarks - vSphere Production and vSphere Recovery

These benchmark tests were performed on the above vSphere environment.

The following benchmarks were taken on different numbers of VPGs and VMs failing over at the same time.

In vSphere environments, the recovery VM is available before promotion of the journal is complete.

# of VPGs	# of VMs in VPG	Total VMs	Total RTO Measured on Zerto 8.0 {MM:SS}
20	1	20	100:50
40	1	40	01:30
60	1	60	01:54
80	1	80	02:43
100	1	100	03:20
1	100	100	04:52

RTO Benchmarks - Hyper-V Production and Hyper-V Recovery

These benchmark tests were performed on the above Hyper-V environment on the production site. The recovery site was Hyper-V.

The following benchmarks were taken on different numbers of VPGs and VMs failing over at the same time.

In Hyper-V environments, the recovery VM is available before promotion of the journal is complete.

# of VPGs	# of VMs in VPG	Total VMs	Total RTO Measured on Zerto 8.0 {MM:SS}
20	1	20	02:10
40	1	40	03:28
60	1	60	07:07
80	1	80	06:16

# of VPGs	# of VMs in VPG	Total VMs	Total RTO Measured on Zerto 8.0 {MM:SS}
100	1	100	07:45
1	100	100	10:13

RTO Benchmarks - VMware Cloud Director Production and VMware Cloud Director Recovery

These benchmark tests were performed on the above VMware Cloud Director (VCD) environment on the production site. The recovery site was also VCD. The following benchmarks were taken on different numbers of VPGs and VMs failing over at the same time. In VCD environments, the recovery VM is available before promotion of the journal is complete.

# of VPGs	# of VMs in VPG	Total VMs	Total RTO Measured on Zerto 8.0 {MM:SS}
20	1	20	04:31
40	1	40	06:42
60	1	60	09:58
80	1	80	12:22
100	1	100	16:16

ZVM GUI Response Times

Alerts generally populate in 5 seconds from the detection time. The GUI is updated in intervals of up to 12 seconds.

The below benchmarks for GUI response times in seconds were measured for typical GUI tasks in the vSphere environment:

	Production Site {seconds}	Recovery Site {seconds}
List VPGs	4	5
General VMs view	3	4
Sites tab	<1	<1

	Production Site {seconds}	Recovery Site {seconds}
VRAs tab	<1	<1
Datastores tab	3	12
Alerts monitor	<1	<1
Events monitor	2	4
VPG performance report	8	7
Outbound protection over time report	2	3-4
Protection over time by site	2	2-4
Recovery report	1-2	2-3
Resources report	not relevant	10-45
Dashboard	4	4
Create VPG	5	4
Edit VPG	4	3

API Response Times

The below table shows measured API response times in the vSphere environment.

API	Production Site Response Time (Seconds)	Recovery Site Response Time (Seconds)
Events	<1	<1
Alerts	<1	<1
VMs	2	3
Tasks	<1	<1
VRAs	<1	<1

API	Production Site Response Time (Seconds)	Recovery Site Response Time (Seconds)
Peer Sites	<1	<1
Hosts	<1	<1
Resourcepools	<1	<1
Datastores	<1	<1
VPGs	2.5	4.5
VPG Checkpoints	<1	<1

Initial Synchronization Benchmarks

In benchmarking tests, the maximum initial sync rate on the above VMware environments is 132 Mbps.

Estimating WAN Bandwidth for VMware and Hyper-V

Note: With Zerto Planner, you can perform the tasks described in this section more efficiently.

Estimating the bandwidth requirements between the protected and recovery sites involves the following:

1. **Collecting data characteristics for protected VMs.** See the following:
 - [vSphere Environments: Collecting Data Characteristics for VMs on page 24](#)
 - [Hyper-V Environments: Collecting Data Characteristics for VMs on page 29](#)
2. **Calculating the estimated bandwidth requirements.** See the following:
 - [Calculating the Estimated Bandwidth Requirement on page 31](#)

Note: When the recovery site is **Amazon Web Services (AWS)**, you estimate the required bandwidth for the protected machines, either in **vSphere** or **Hyper-V** as described below.

vSphere Environments: Collecting Data Characteristics for VMs

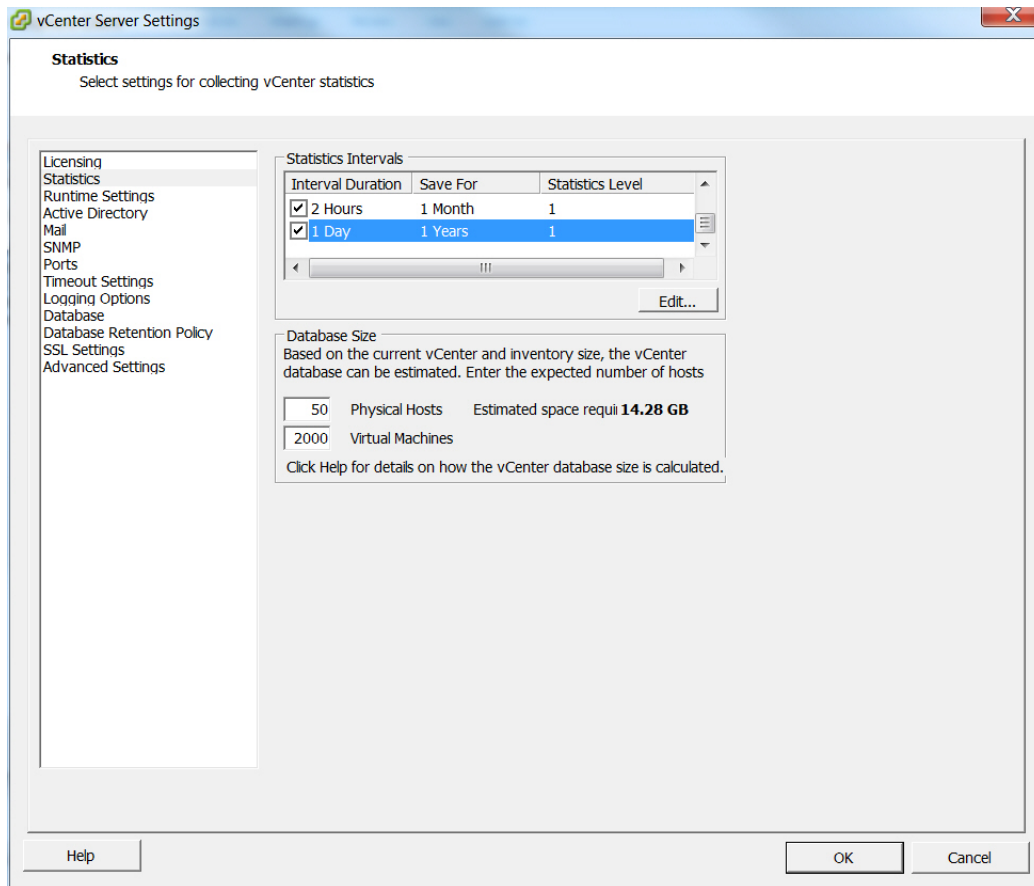
Before you can collect the required data, you must first enable data collection in vCenter Server.

- [Enabling vCenter Server Data Collection on page 24](#)
- [Collecting Data Characteristics for VMs on page 26](#)

Enabling vCenter Server Data Collection

> To enable vCenter Server data collection:

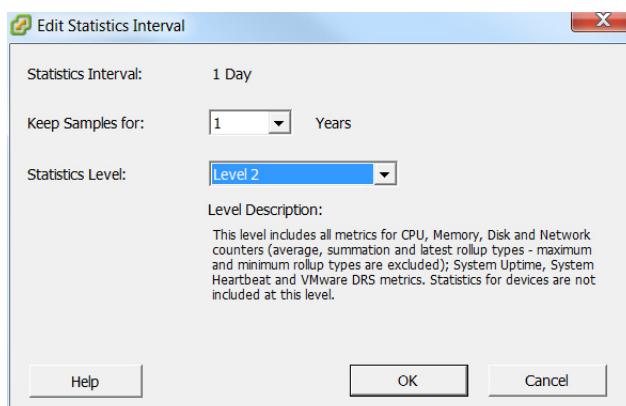
1. Connect to the vCenter Server.
2. In the Administration menu item, select **vCenter Server Settings**.
The vCenter Server Settings dialog is displayed.
3. Select **Statistics**.



4. Make sure that the **Statistics Level** value for all interval durations up to and including the one day duration is at least 2.

If any of the durations have a value less than 2, do the following, starting with the smallest interval:

- a. Select the interval and click **Edit**.
- b. Change **Statistics Level** to **Level 2**.
- c. Click **OK**.



5. Repeat [step 4](#) for all the values up to and including the 1 day interval duration.

6. Click **OK** and wait for at least a day before using the aggregate usage data.

Collecting Data Characteristics for VMs

You can collect data characteristics for the virtual machines in a VPG in one of the following ways:

- Via **vSphere Client** console performance statistics. See [Collecting Data Characteristics for VMs via vSphere Client Console Performance Statistics on page 26](#).
- Or -
- By running a **script** to collect the data characteristics. See [Running a Script to Collect the Data Characteristics on page 28](#).

Note: The script samples supplied with the download, require **vSphere PowerCLI** and permissions to access the vCenter Server using the script.

- Or -

- By using operating system performance monitors, such as the Microsoft Performance Monitor utility for Windows operating systems, or the `iostat` command for Linux operating systems.

Collect data for a **minimum of one day**. Collecting this information impacts on performance and therefore the collection period should be long enough to gather a true representation of usage, but not too long.

Use one of the following procedures:

- [Collecting Data Characteristics for VMs via vSphere Client Console Performance Statistics on page 26](#) uses a timeframe of **one** day.
- [Running a Script to Collect the Data Characteristics on page 28](#) uses a timeframe of **seven** days.

Note: When running vCenter Server versions **before version 5.x**, if any of the virtual machines use NFS storage, metrics for the NFS storage are not generated by the vCenter Server.

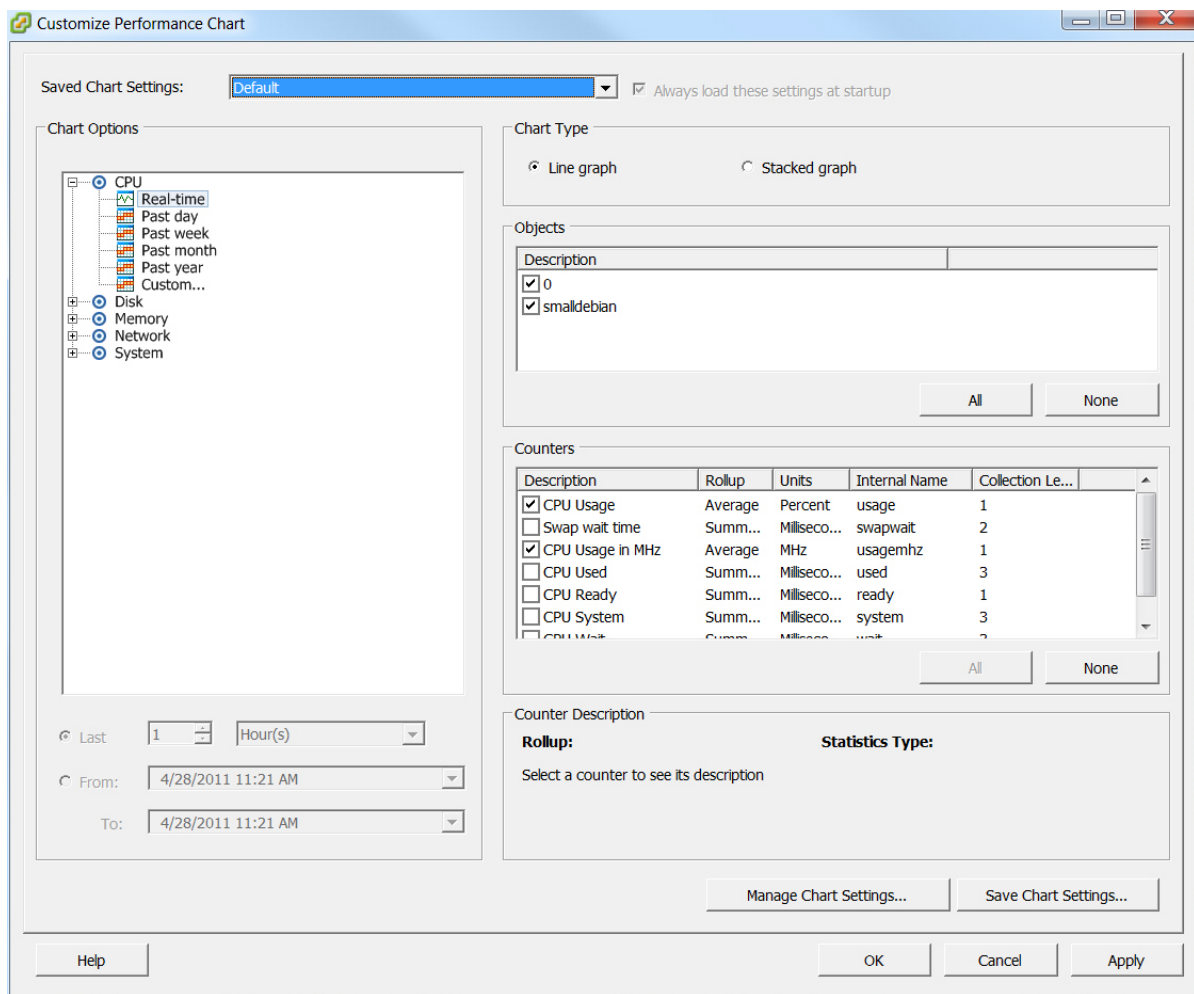
Collecting Data Characteristics for VMs via vSphere Client Console Performance Statistics

Use the following procedure to collect data characteristics for the VMs via the vSphere Client console performance statistics.

> To collect data characteristics for the VMs via the vSphere Client console performance statistics:

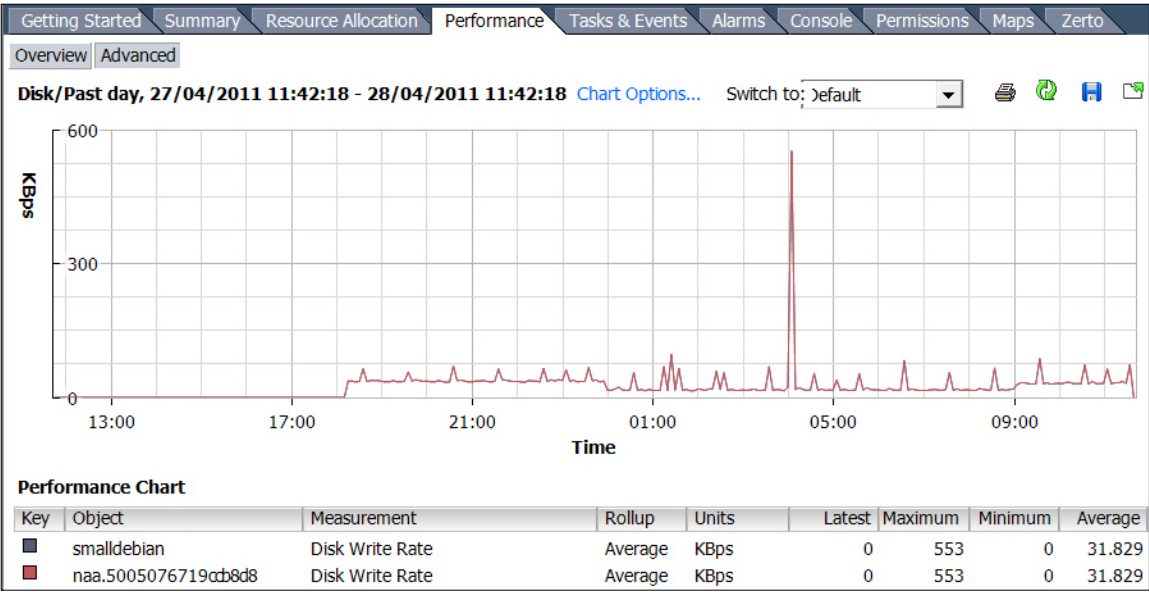
1. In the vSphere Client console select the VM and open the **Performance** tab.
2. Click **Advanced**.
3. Click the **Charts Options** link.

The Customize Performance Chart dialog is displayed.



- In the area **Chart Options**, drill-down in **Disk** and select **Past day**.
- In the area **Counters**, click **None** to clear all the selections and then select **Disk Write Rate or Write Rate**.
- Click **OK**.

A chart like the following is generated:



Use the chart for the average write rate of the VM.

Running a Script to Collect the Data Characteristics

Use the following procedure to run a script to collect the data characteristics.

Note: The following script and the samples supplied with the download, require vSphere PowerCLI and permissions to access the vCenter Server using the script.

> **To collect data characteristics for the VMs via a script:**

1. Run a script like the following:

```
$report = @()
Get-VM | %{
    $stats = Get-Stat -Entity $ -Stat disk.write.average -Start
(Get-Date).adddays(-7) -ErrorAction SilentlyContinue
    if($stats){
        $statsGrouped = $stats | Group-Object -Property MetricId
        $row = "" | Select Name, WriteAvgKBps, WriteAvgMBps
        $row.Name = $_.Name
        $row.WriteAvgKBps = ($statsGrouped |
            where {$_.Name -eq "disk.write.average"} |
            %{$_.Group | Measure-Object -Property Value -
Average}).Average
        $row.WriteAvgMBps = $row.WriteAvgKBps/1024
        $row.WriteAvgKBps = "{0:N2}" -f $row.WriteAvgKbps
        $row.WriteAvgMBps = "{0:N2}" -f $row.WriteAvgMBps
        $report += $row
    }
}
$report | Export-Csv "C:\ZertoOutput.csv"
```

2. If you want a value other than seven days, change the value of the **adddays()** function. For example to collect data for three days, use **adddays(-3)**.
3. Use the resulting file, **C:\ZertoOutput.csv**, for the average write rate of the VM.

Note: Versions of this script are included in the download with this document.

4. Continue with [Calculating the Estimated Bandwidth Requirement on page 31](#).

Hyper-V Environments: Collecting Data Characteristics for VMs

You can collect data characteristics for the virtual machines in a VPG in one of the following ways:

- By using operating system performance monitors, such as the Microsoft Performance Monitor utility for Windows operating systems or the `iostat` command for Linux operating systems.
- By using Windows PowerShell in Windows Server 2012 to collect network utilization (and other information). When using metering ACLs, you can measure the total network traffic sent and received by a virtual machine. To collect performance characteristics for the virtual machines in a VPG, using PowerShell, do the following:
 - Turn on resource metering for the relevant virtual machines, if it is not already enabled.
 - Adjust the collection frequency, if necessary.
 - Collect the relevant statistics.

Zerto recommends that you collect data for a minimum of one day. When you have enough statistics, you may want to turn off resource metering since data collection can impact performance.

- [Turning on Resource Metering on page 30](#)
- [Adjusting the Collection Frequency on page 30](#)
- [Collecting and Viewing the Relevant Statistics on page 30](#)
- [Turning off Resource Metering on page 31](#)

Turning on Resource Metering

By default, resource metering is not enabled. To turn on resource metering for one virtual machine, enter the following PowerShell command:

Get-VM <VM-name> | Enable-VMResourceMetering

To turn on monitoring for all virtual machines on a server at one time, enter the following PowerShell command:

Get-VM | Enable-VMResourceMetering

Once you enable resource metering, Hyper-V begins to collect data. You can reset metering at any time, which discards the data that has been collected up to that point.

If resource metering is enabled but no NetworkAdapterAcls are configured, Hyper-V configures them to measure total network traffic. To measure network traffic through an IP range, configure the NetworkAdapterAcls for the IP range before running **Enable-VMResourceMetering**.

Adjusting the Collection Frequency

By default, the collection frequency is once every hour. You can change the collection frequency but understand that data collection can impact performance. To change the collection frequency, enter the following command:

Set-VMHost -ComputerName <host-server-name> -ResourceMeteringSaveInterval <HH:MM:SS>

The collection frequency is always set at the host server level. You cannot adjust the collection frequency per virtual machine.

For example, if you enter 01:30:00, resource consumption will be collected every hour and a half.

Collecting and Viewing the Relevant Statistics

To view resource usage for one virtual machine, enter the following command:

Get-VM <VM-name> | Measure-VM

Resource metering data can be displayed for all of the virtual machines that are running on a host. To see data for all of the virtual machines on a host, enter the following command:

Get-VM | Measure-VM

You can configure PowerShell to display only certain statistics. To do this, you must know the object names that PowerShell assigns to each statistic. You can see the object names by entering the following command:

Get-VM | Measure-VM | Select-Object *

For example, when working with Zerto, you are interested in network traffic. To list the network traffic for each virtual machine, enter the following command:

Get-VM | Measure-VM | Select-Object VMName, NetworkMeteredTrafficReport

You can use VM Network Adapter ACLs to measure network activity to and from a specific network. For example, to meter network traffic for a special subnet or IP address:

Add-VMNetworkAdapterAcl -VMName <VM-name> -Action Meter -RemoteIPAddress 10.10.0.0/16 -Direction Outbound

Turning off Resource Metering

To disable the collection of performance statistics, enter the following PowerShell command:

Disable-VMResourceMetering -VMName <VM-name>

Continue with [Calculating the Estimated Bandwidth Requirement on page 31](#).

Calculating the Estimated Bandwidth Requirement

Note: The Zerto WAN Sizing Estimator is included in the download with this document.

Use the **average write rate** for the virtual machines in a VPG in the Zerto WAN Sizing Estimator to estimate the minimum bandwidth required.

For each VM you also must decide whether **compression** will be enabled for the VM, based on the data characteristics.

Use one of the following procedures:

- [Estimating sizing using the Zerto WAN Sizing Estimator on page 31](#)
- [Estimating sizing without using the Zerto WAN Sizing Estimator on page 32](#)

Estimating sizing using the Zerto WAN Sizing Estimator

> To estimate sizing using the Zerto WAN Sizing Estimator:

1. Open the **Zerto WAN Sizing Estimator**.
2. Enter the following information in the **VM data** sheet:
 - The **VM** name.

- The **Write KB/s** data, based on the statistics gathered in the previous task. Use a period for the decimal mark.
- Define whether **compression** is enabled for this VM: Select **Yes** or **No**.
- The **application data** characteristics: Select **Compressed** or **Compressible**.

The Zerto WAN Sizing Estimator colors the cell red if you decide to employ compression on compressible data and orange if you decide to avoid compression for compressible data.

3. The Zerto WAN Sizing Estimator estimates the **total bandwidth** needed for your deployment, using a **minimum** value of **5 Mb/sec**.

The estimation is displayed on the top of each page of the Zerto WAN Sizing Estimator.

To estimate the **WAN sizing** required **without** using the Zerto WAN Sizing Estimator, using the following procedure.

Estimating sizing without using the Zerto WAN Sizing Estimator

> To estimate sizing without using the Zerto WAN Sizing Estimator:

1. For **each** virtual machine in the VPG multiply the KB/sec (which is based on the statistics gathered) by 8, and divide the result by 1024 to provide an answer in Mb/sec.
 - If **compression is enabled** for the VM and the data is compressible, divide this result by **2**.
2. Sum the results of [step 1](#).

WAN Mb/sec = SUM(KB/sec * (8/1024/(1 or 2 if compressible data that will be compressed)))
--

The result is an estimate of the required Mb/sec for the WAN.

Note: If the result is less than 5 Mb/sec, see [Zerto Scale and Benchmarking Guidelines](#) for minimum bandwidth requirements.

Zerto enhances the Zerto IT Resilience Platform by converging disaster recovery and backup to deliver continuous availability within a simple, scalable platform. Zerto delivers enhanced analytics, platform improvements and cloud performance upgrades required in the future of IT resilience.

Learn more at Zerto.com.

For assistance using Zerto's Solution, contact: [@Zerto Support](https://twitter.com/ZertoSupport).

© 2020 Zerto Ltd. All rights reserved.